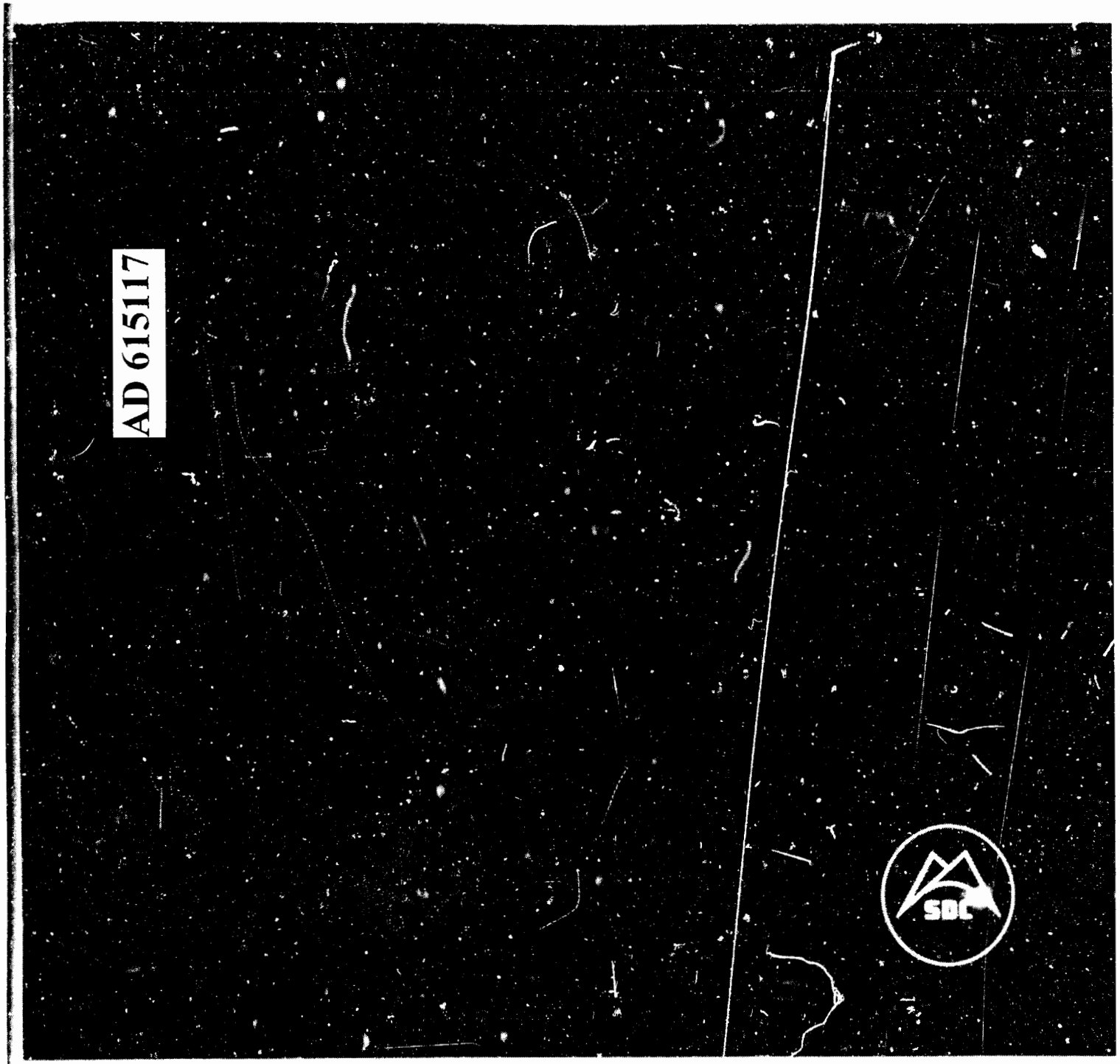

SP-2059

RESEARCH IN THE NAGEMENT OF COMPUTER PROGRAMMING

SP-2059

SP *a professional paper*

RESEARCH IN THE MANAGEMENT OF COMPUTER PROGRAMMING

by

George F. Weinwurm

1 May 1965

SYSTEM DEVELOPMENT CORPORATION
2500 COLORADO AVE.
SANTA MONICA
CALIFORNIA

This material draws heavily from several SDC documents by L. Farr, B. Nanus, and especially V. LaBolle dealing with the same general topic.

(Page 2 blank)

ABSTRACT

Computer programming is a relatively new and rapidly growing segment of Automatic Data Processing and is an area largely unexplored from the standpoint of management science. The management measures and standards that are common to many, more mature industries are nearly non-existent in computer programming.

The Paper describes the work that has been underway at the System Development Corporation for almost three years directed toward the advancement of the state-of-the-art of computer programming management.

COMPUTER PROGRAMMING: BIG AND GROWING

The acceptance of Automatic Data Processing as a normal part of the business, military, and scientific scene has come about so rapidly and pervasively that even the optimistic predictions made in recent years are now thought of as too conservative. The frenzied pace of technical competition and discovery shows no sign of slackening, and the task of forecasting the development of the computer field will probably continue to be as hazardous an occupation in the future as it has been in the past.

Although their projections of magnitude may vary, however, experts are in complete agreement on one point: the production and installation of computer hardware and the associated programs has become a very big busines, indeed, and seems destined to become larger still. Estimates from various industry and Government sources are that computer programming expenditures alone will range between $3 and $7 billion annually by 1970 (1).

It is only natural that the prospect of spending such enormous amounts of money for computer systems has focused attention on the managerial questions of performance, quality, and effectiveness. This is especially the case with the Federal Government, which is by far the largest single user of Automatic Data Processing equipment and services. Two significant studies along these lines have been issued recently: "Review of Problems Relating to Management and Administration of Electronic Data Processing Systems in the Federal Government," by the Comptroller General of the United States (1); and the so-called "Clewlow report" on "The Management of Automatic Data Processing in the Federal Government," (2) which was prepared by the Bureau of the Budget. Both of these deal with suggested management criteria for evaluating the potential acquisition and use of computer systems and the associated programs.

It is generally recognized, however, that these Reports represent more of a first step and identification of areas for future work rather than a definitive solution. A great deal needs to be done before the present profusion of computer terminology and techniques, which have been likened by General David Sarnoff in a recent speech (3) to a "technological Tower of Babel" have been analyzed and synthesized into measures and standards that are useful to a manager in the performance of his day-to-day responsibilities.

THE MANAGER'S DILEMMA

One of a manager's primary tasks is to review the estimates of cost and schedule within which a project will operate, and to balance these against product performance over a period of time. In modern parlance, decisions of this sort are often identified by such labels as "cost/effectiveness."

Unfortunately, the computer programming field is so young and so dynamic that no generally accepted standards and techniques have been developed that permit the manager to make predictions and comparisons. Of course, judgements are made as they are needed, but confidence in them is low and the difference between before-the-fact estimates of cost and schedule and after-the-fact history is often depressingly large.

The difference between the managerial state-of-the-art in computer programming and that which prevails in more mature industries can be perceived most clearly from the point of view of the potential buyer of a computer program and, a potential buyer of, say, an automobile.

First of all, an automobile is a tangible piece of equipment whose qualitative attributes can be assessed rather easily. A computer program, on the other hand, is delivered as a set of documents, cards, tapes, and listings, that represent an operational entity that cannot be seen, heard, smelled, or kicked.

Furthermore, the potential buyer of an automobile can make use of the evaluations of popular journals in the field, descriptive literature from the manufacturers, or subjective reactions of friends who "own one." Most of these facts and opinions will be phrased in terms of measures common to automobiles, such as horsepower, turning radius, comfort, style, safety, optional accessories, freedom from repair, and so on.

One who is considering the purchase of a computer program has a much more difficult task in establishing any reasonable criteria on which to make comparisons. Few programs are ready-made; most are tailored to the needs of the user. There are almost no standards for comparing the characteristics of programs with their expected performance. Usually, the design of a computer program is based upon a description of the job to be done, but most often the characteristics are highly qualitative rather than quantitative.

Now, suppose that these contrasts between products are seen through the eyes of a production manager rather than a potential buyer. The manager in the automobile factory benefits immediately from the nomenclature and procedures that are standard throughout the automotive industry. There are techniques available to predict and measure the performance of the product, on both

a unit-by-unit and a sampling basis; standards against which the output of both men and machines can be gauged; and an abundance of historical material in terms of which the efficiency of alternatives can be estimated.

Because computer programming projects tend to be "one-shot," or at least are usually thought of in this way, the manager of such an activity is faced with an almost total lack of standard measures for the performance or quality of products or tasks, or predictive techniques for the costs and manpower that will be required. In addition, there is no generally available body of historical data on which to rely in selecting alternative courses of action. Clearly, the computer programming manager would benefit immensely if the kinds of management measures, standards, and techniques that are common to many other industries, such as the automotive, could also be developed for the field of computer programming.

THE PROBLEM SUMMARIZED

Specifically, the evolution of such managerial aids for computer programming has been hindered by several obstacles:

- The most apparent is simply the youth and technical turbulence of the computer programming field. So much has come about so rapidly that comparatively little time has been spent on synthesis or management research.

- Beyond this, there is a general lack of agreement on the terminology in use throughout the computer programming field. While there are many glossaries, such as those prepared by the Association for Computing Machinery (ACM) and the Bureau of the Budget, they are not very widely used, and most organizations evolve their own set of working definitions.

- Little attention has been given to the definition of attributes that characterize the nature or quality of a computer program as a product. For example, programmers use such terms as "maintainability," "tightness of coding," and "flexibility," but there seem to be no widely accepted criteria by which similar programs could be compared in terms of these factors.

- Present cost collection criteria seem to be designed primarily for legal accounting purposes. For this reason, the historical data that remain after a programming project has been completed are not readily adaptable to analysis in terms of managerial planning and control. Generally speaking, non-cost historical information is not kept in any organized and cohesive fashion at all. What records

remain are not usually comparable from one organization to another and, often, not even within different portions of the same organization.

- Many of the design, schedule, and resource constraints that impinge upon the computer programming process in some way are not well enough understood to be susceptible to quantification. The interrelationships between the steps in the program production process itself are not well understood and no general agreement exists on how the diverse variations should be combined to obtain conclusions that are meaningful in terms of the program end-product.

RESEARCH IN COMPUTER PROGRAMMING MANAGEMENT

The System Development Corporation has been engaged in a research project focused on these problems of computer programming management for nearly three years, first under the sponsorship of the Advanced Research Projects Agency of the Assistant Secretary of Defense for Research and Engineering, and later with the support of the Air Force Systems Command and the Office of Naval Research.

The aim of the Project has been to develop techniques that will permit computer programming managers to reduce lead times and costs and improve the quality of the product; or, in a more general sense, to "identify, extend and unify scientific knowledge pertaining to (computer programming) management (4)."

The work itself has been structured in terms of the problems that were deliniated previously. During the past three years, significant progress has been made in several areas:

- The computer program development process has been thoroughly dissected and analyzed, and a number of common steps defined. Among the products of this effort have been rather detailed Planning Guides for computer programming managers.

- A substantial body of historical data, both quantitative and qualitative, has been collected for a wide variety of completed computer programming projects. A detailed Questionnaire has been developed and refined for this purpose. This data base is being expanded continually, as current programming projects are sampled.

- More than one hundred resource-cost factors that affect computer programming activities have been identified. Multiple-regression and other statistical techniques have been applied to the data base to narrow the range of resource-cost factors that are truly important, and around which incremental cost collection systems can be constructed.

- Some preliminary work has been done toward the development of management standards in terms of which the performance and quality of computer programming products can be measured.

At the present time, the type of help that the Project is able to offer the programming manager is largely qualitative, as represented by the Planning Guides, rather than quantitative. However, some general conclusions along the latter lines are beginning to be available from the statistical data base analysis. Figures 1 and 2 illustrate the relationships between three resource-cost factors--new instructions (5), man months, and computer hours--in terms of a band of 67-percent confidence. Although the variance is still rather large, comparative historical data of this type have already proved their usefulness as a base against which intuitive estimations can be checked.

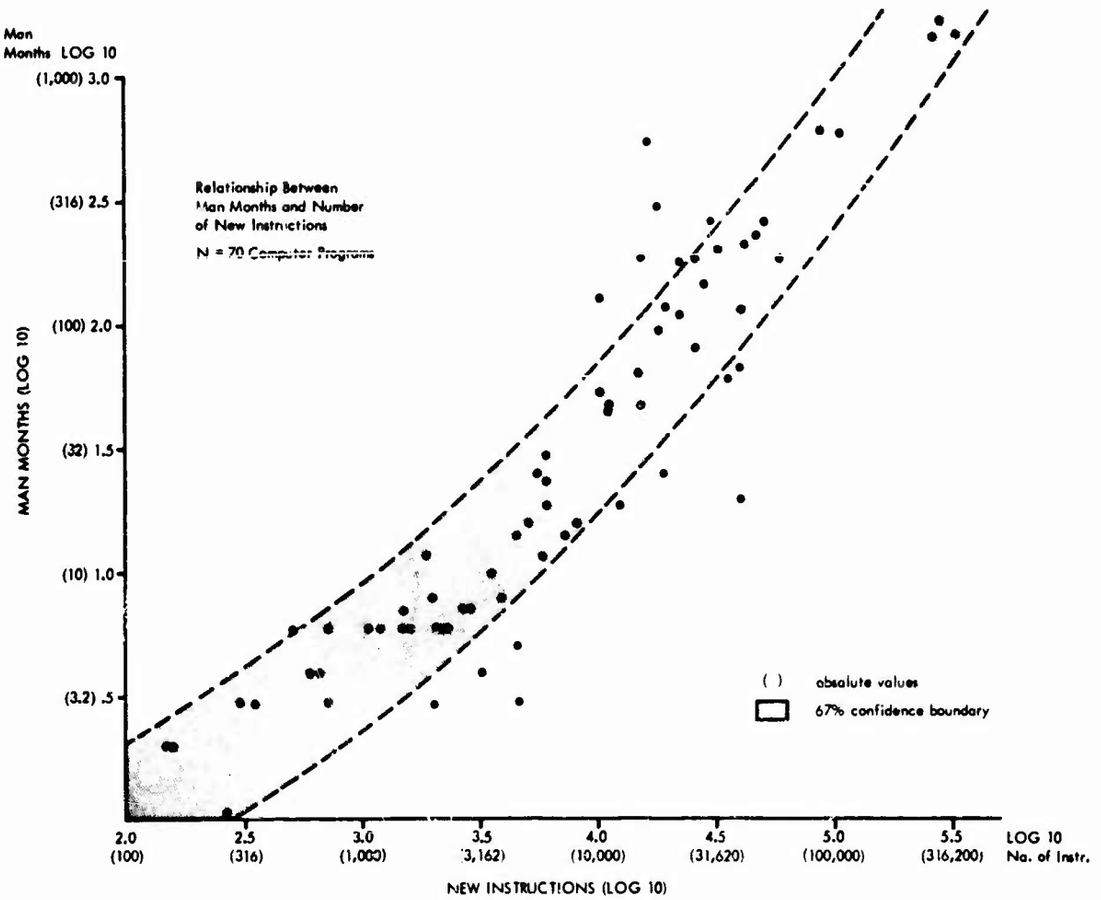
Man
Months LOG 10
(1,000) 3.0
(316) 2.5
(100) 2.0
(32) 1.5
(10) 1.0
(3.2) .5
MAN MONTHS (LOG 10)
Relationship Between
Man Months and Number
of New Instructions
N = 70 Computer Programs
( ) absolute values
67% confidence boundary
2.0 (100)
2.5 (316)
3.0 (1,000)
3.5 (3,162)
4.0 (10,000)
4.5 (31,620)
5.0 (100,000)
5.5 (316,200)
LOG 10 No. of Instr.
NEW INSTRUCTIONS (LOG 10)

Figure 1

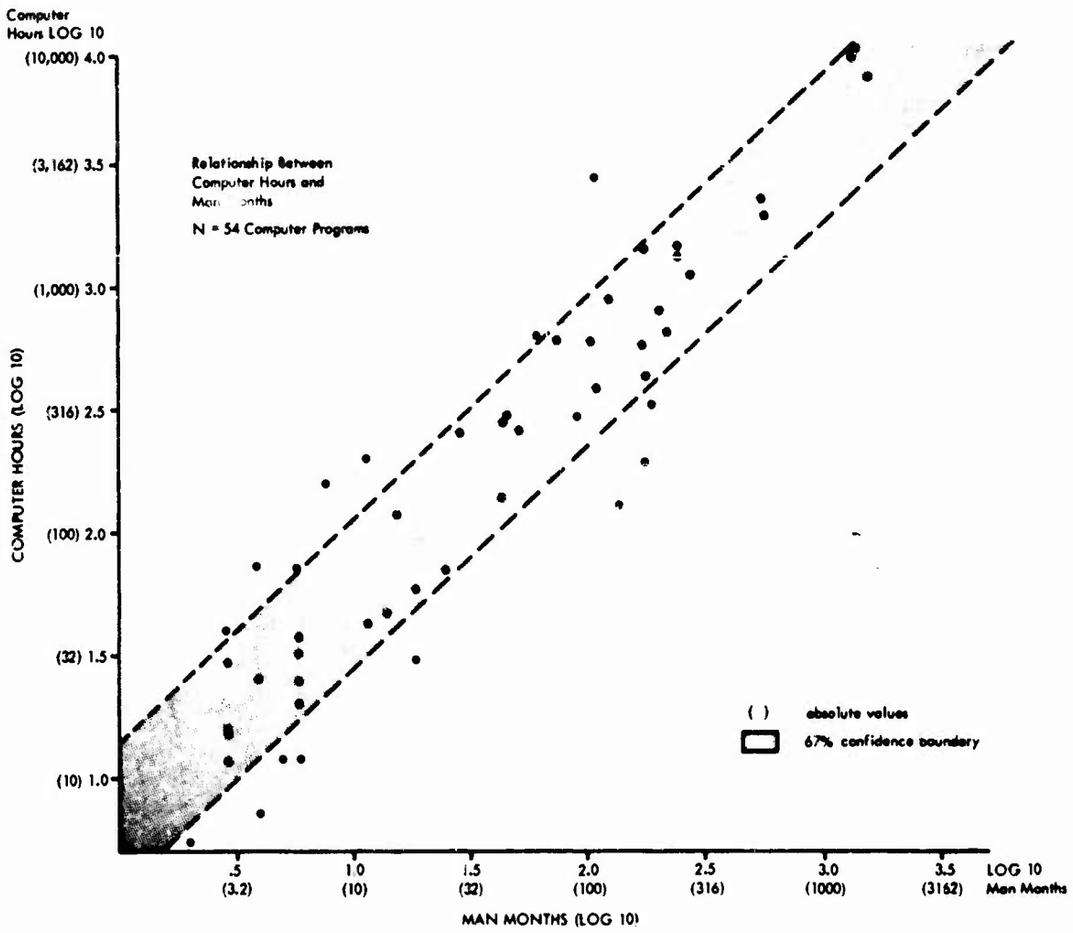
Computer Hours LOG 10
(10,000) 4.0
(3,162) 3.5
(1,000) 3.0
(316) 2.5
(100) 2.0
(32) 1.5
(10) 1.0
COMPUTER HOURS (LOG 10)
Relationship Between Computer Hours and Man Months
N = 54 Computer Programs
( ) absolute values
67% confidence boundary
.5 (3.2)
1.0 (10)
1.5 (32)
2.0 (100)
2.5 (316)
3.0 (1000)
3.5 (3162)
LOG 10 Man Months
MAN MONTHS (LOG 10)

Figure 2

RESEARCH PLANS FOR THE FUTURE

The main emphasis at the moment is on the expansion and refinement of the questionnaire-acquired data base so that the statistical analysis will become increasingly meaningful and reliable. The Project has been granted a Public Reports Approval Number by the Bureau of the Budget and the Air Force Systems Command to sample programming projects underway or recently completed at a number of industrial and military organizations. The analysis of this data will continue through the balance of 1965 and into early 1966, when a comprehensive report will be published and made available to all participants.

This particular line of research has recently been focused on the problems of predicting, collecting, and controlling computer programming costs. The intent is to define for the manager the resource-cost factors that are of prime importance for these purposes. The resulting factors will likely be somewhat different from those in terms of which costs are being collected generally as a part of existing accounting systems, but will have demonstrated a high degree of meaning in terms of the computer programming manager's special needs.

COMMENTS AND REFERENCES

1. In a 1963 article ("A Profile of the Programmer," Industrial Relations News, August, 1963), Deutch and Shea estimated the 1970 requirement for programmers to be in the neighborhood of 200,000. About a year later, Brandon ("The Computer Personnel Revolution," Computers and Automation, August, 1964) projected a need for about 145,000 programmers and 90,000 system analysts in 1970, based on the computer installations anticipated through that time period. Doubling an average salary of $10,000 (which is estimated for 1970 based on an extensive annual SDC National Salary Survey for digital computing personnel) to obtain gross costs, and assuming a demand for 200,000 programmers and system analysts, this yields annual expenditure of $4 billion for computer programming alone by 1970.

   Such an estimate is comparable to those contained in the 1963 "Survey and Study of the Computer Field" by the Investment Bankers Association of America (contained in Use of Electronic Data Processing Equipment in the Federal Government, October 16, 1963, published hearings before the Committee on Post Office and Civil Service, House of Representatives, 88th Congress, First Session) and the recent Review of Problems Relating to Management and Administration of Electronic Data Processing Systems in the Federal Government (The Comptroller General of the United States, August, 1964). Assuming the rule of thumb that investments in program development usually equal or exceed those in computer hardware, the Investment Bankers project an annual expenditure range of from $4 to $7 billion by 1970, while the Comptroller General's Report is somewhat lower at $3 billion.

2. Committee on Government Operations, U. S. Senate, Report to the President on the Management of Automatic Data Processing in the Federal Government, March, 1965. This Report was prepared for the Bureau of the Budget by a Project Staff including Carl W. Clewlow.

3. Presented at the Fall Joint Computer Conference, San Francisco, California, on October 27, 1964.

4. The quotation is the theme of TIMS (The Institute of Management Science).

5. Note that the x-axis for Figure 1 is in terms of new instructions rather than total instructions.